



University of Pennsylvania Working Papers in Linguistics

Volume 2

Issue 2 *Proceedings of the 19th Annual Penn
Linguistics Colloquium*

Article 13

1-1-1995

Certain Aspects of Cliticization in Polish

Adam Szczegielniak

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/pwpl/vol2/iss2/13>
For more information, please contact libraryrepository@pobox.upenn.edu.

Certain Aspects of Cliticization in Polish

CERTAIN ASPECTS OF CLITICIZATION IN POLISH

Adam Szczegielniak
Harvard University
SZCZEG@FAS.HARVARD.EDU

1. Introduction, the Data, and Historical Background¹

The main goal of this paper is to present an account of certain forms of cliticization in Polish. I will try to show that the fact that certain clitics undergo phonological processes typically assumed to be lexical does not exclude the possibility of them being generated in the syntax, hence this paper is also an attempt to provide a framework of the interaction between different modules of grammar. Clitics are a good testing ground since their behaviour is distinct in different levels of grammar, especially syntax and phonology.

Polish preterite endings and Polish *by* 'would' can be considered distinct from affixes and independent words either syntactically, or phonologically, and hence should be considered to be clitics. The examples presented here are considered by the majority of linguists to be clitics (see Booij & Rubach (1987), Dogil (1984), Aguado & Dogil (1989), Spencer (1991), Zwicky (1977)). This paper will mainly concentrate on the following clitics:

- (1) - *m* : 1sg. preterite ending, eg.. *zab+i+l+e+m* 'I killed'
- *ś* : 2sg. preterite ending, eg.. *zab+i+l+e+ś* 'you killed'
- *\$my* [*\$mɨ*]: 1pl. preterite ending, eg.. *zab+i+l+i+\$my* 'we killed'
- *\$cie* [*\$tɕe*]: 2pl. preterite ending, eg.. *zab+i+l+i+\$cie* 'you killed'
- *by* [*bɨ*]: 'would', eg.. *zab+i+l+by* 'he would kill'

Following Klemensiewicz and Urbańczyk (1955:37), Booij & Rubach (1987:41) indicate that the preterite clitics derive historically from the forms of the auxiliary 'to be':

Singular: *jeśm* -> *m*; *jeś* -> *ś*. Plural: *jeśmy* -> *\$my*; *jeście* -> *\$cie*.

For Booij & Rubach, the lexicalisation of the above forms is proof that they are derived in the lexicon. However, if we assume that they are still generated in the syntax, and yet have the phonological properties usually assigned to affixes, we are able to indicate that the historical development of these forms seems to indicate that they are evolving into affixes.

2. The Distribution of Clitics

The preterite clitics most often attach to verbs, however, they may also attach to other constituents that precede the verb. Consider the following examples (following Booij & Rubach (1987: 34), the clitic is in italics, a plus denotes affixation.):

- (2) (a) PRONOUNS:
ja to rob+i+l+e+m = *ja+m to rob+i+l* = *ja to+m rob+i+l* 'I did this'
- (b) INTERROGATIVE PRONOUNS AND PARTICLES:
co ja rob+i+l+e+m = *co +m ja robił* 'what did I do'
czy ja tam by+l+e+m = *czy+m ja tam by+l* 'was I there'

¹ I would like to thank all the anonymous reviewers of my abstract submitted to the 19 Penn Linguistics Colloquium, as well as to the audience of the Fifth Conference on Generative Grammar in La Coruña. Special thanks also to Dorota Wojtaś and Krystyna and Wiesław Szczegielniak. Needless to say all errors are mine. In the paper herein, I have not taken into account recent work by Borsley and Rivero (1994), as well as the recent developments in Chomsky's Minimalist Programme.

(c) CONJUNCTIONS AND COMPLEMENTIZERS:

nie kazali mi ale rob+i+l+e+m 'they didn't tell me but I did (it)'

= nie kazali ale+m rob+i+l

myslał, że tam by+l+e+m

'he thought I was there'

= myślał, że+m tam by+l

(d) ADVERBS:

szybk+o to rob+i+l+e+s = szybk+o+s to rob+i+l 'you did it fast'

(e) NOUNS:

dla Stefan+a to rob+i+l+e+s 'you did (it) for Stefan'

= dla Stefan+a+s to rob+i+l

The above examples, showing how freely the preterite clitics can attach to different words, clearly indicate why we cannot consider the preterite clitics to be affixes. The distribution of Polish *by* is similar. Consider the distribution of *by* within a clause (this holds also for the preterite endings):

- (3) (a) *On zrob+i+l+by to* 'he would have done that'
HE DID+CL THAT
- (b) *On by zrob+i+l to* 'he would have done that'
HE CL DID THAT
- (c) *On to by zrob+i+l* 'he would have done that'
HE THAT CL DID
- * (d) *On zrob+i+l to by* 'he would have done that'
HE DID THAT CL
- (e) *Teraz by on zrob+i+l to* 'now he would have done that'
NOW CL HE DID THAT
- (f) *Coby on zrob+i+l* 'what he would have done'
WHAT CL HE DID
- * (g) *By on zrob+i+l to* 'he would have done that'
CL HE DID THAT

Examples (3d, g) clearly show that there are certain specific syntactic positions within the clause where *by* cannot move to (sentence initial and right of V).

Szczegielniak (1991) argues against Booij & Rubach's (1987) proposal that clitics find their phonological hosts in the Lexicon. Following Aguado & Dogil (1989) and Szczegielniak (1991)², I will claim that the preterite endings are generated in the INFLAGR node, whereas *by* is either generated in INFL or COMP, depending on the selectional properties of the main clause verb. Consider the following sentences:

- (4) (a) *On chciał, by Andrzej to zrobił* 'he wanted Andrew to do it'
he want+past CL Andrew it do+past
- * (b) *On chciał, (że) Andrzej by to zrobił*
he want+past (that) Andrew it do+past
- (c) *On wiedział, że Andrzej by to zrobił*
he know+past that Andrew CL it do+past
- 'he knew that Andrew would do it'
- * (d) *On wiedział, żeby Andrzej to zrobił*
he know+past that+CL Andrew it do+past
- * (e) *On wiedział, że Andrzej zrobił toby*
he know+past that Andrew do+past it+CL

From (4a-b) we should conclude that *by* obligatorily must be present in the COMP position, and, what is more, it cannot move anywhere within the subordinate clause.³

² The two major arguments against the generation of the preterite clitics in the lexicon are: 1. Over-generation in the lexicon, since every word in Polish would have to be listed with the clitic and without it; 2. The need to postulate syntactic filters making use of phonological information in order to exclude constructions like: *X.....Y+CL1.....Z+CL1....V.

³ I consider *żeby*, *aby*, *by* to be phonetic variants of the same unit.

However, if we take a different main clause verb, examples (4c-d), the picture is exactly opposite. It has to be concluded that the verb *chcieć* 'to want' selects *by* in COMP. The verb *wiedzieć* 'to know,' on the other hand, selects only the complementizer *że*, and *by* is generated in INFL. The INFL generated clitic has freedom to move within the subordinate clause because it is selectionally independent of the main clause verb. Example (4e) shows that the clitic cannot attach to elements placed after the verb, however, it can attach to the same elements (object *to*) if they are moved in front of the verb. *Wiedzieć* and *Chcieć* form two categories of verbs in Polish, as far as the selection of a complementizer is concerned. Interesting is the example of the verb *powiedzieć*. When it takes *by* in COMP its meaning is 'to order,' when *by* is selectionally independent, the meaning of *powiedzieć* is 'to speak.' The behaviour of *powiedzieć* indicates that the semantic properties of the main clause verbs somehow are interrelated with the kind of complementizer it selects.

Polish preterite endings show similar distribution to that of the INFL generated by. Consider:

- (5) (a) *On wiedział, że myśmy to zrobili* 'he knew that we did it'
 he know+past that we+CL it do+past
 * (b) *On wiedział, że my tośmy zrobili*
 he know+past that we it+CL do+past
 (c) *On wiedział że my to zrobiliśmy*
 he know+past that we it do+past+CL
 (d) *On wiedział, żeśmy my to zrobili*
 he know+past that+CL we it do+past
 (e) *On wiedział, że my byśmy to zrobili*
 he know+past that we+CL+CL it do+past
 * (f) *On wiedział, że my by to zrobiliśmy*
 he know+past that we CL it do+past+CL
 * (g) *On wiedział, żebyśmy to zrobili*
 he know+past that+CL+CL we it do+past
 * (h) *On wiedział, że my zrobili tośmy*
 he know+past that we do+past+CL
 (i) *On chciał, (że) byśmy my to zrobili* 'he wanted us to do it'
 he want+past (that)+CL+CL we it do+past
 * (j) *On chciał, (że) my byśmy to zrobili*
 he want+past (that) we CL+CL it do+past
 * (k) *On chciał, (że)by myśmy to zrobili*
 he want+past (that) we+CL+CL it do+past
 * (l) *On chciał, (że)by my tośmy zrobili*
 he want+past (that)+CL we it+CL do+past
 * (m) *On chciał, (że)by my to zrobiliśmy*
 he want+past (that)+CL we it do+past+CL

From the above data, it seems that the distribution of *śmy* (and the other preterite endings) generally does not differ from that of *by* generated in INFL. Both can only appear to the left of the subordinate clause verb⁴, compare (4e) with (5f). The distribution of these clitics is similar but not identical. The relevant examples are (4d) and (5d). In the case of (4d) the clitic *by* cannot appear in the COMP position; (5d) shows that this is not true of the preterite clitics. However, I will claim that the reason for such a distribution of *by* is not syntactic but pragmatic. Two factors have to be taken into account: (a) in example (4d) *by* is in the same position as in (3a), where it is selectionally dependent on the main clause verb. There is a possibility that this position in (4d) makes the clitic's status ambiguous, (b) clitic placement to elements other than the verb is motivated pragmatically - emphasis marking. It

⁴Contrary to Booij & Rubach (1987), I assume that both preterite endings and *by* cannot appear sentence initially. Hence, **By oni poszli do kina* 'they would have gone to the cinema', or **śmy my poszli do kina* 'we went to the cinema', are considered to be ill-formed (J. Rubach, personal communication).

is possible that *by* cannot be an emphasis marker in COMP, if not selected by the main clause verb.

The ungrammaticality of (5f) and (5k-m) is caused by the fact that the construction *by*+preterite clitic is obligatory. It is also interesting to note that in (5i-m) the *by*+Clitic construction retains the properties of the *by* generated in COMP. This proves the assumption that the features of the left clitic percolate upwards, and that *by* is the head of the whole construction.

3. Possible Hosts for Cliticization

In 5b *to* 'it' cannot serve as a host for *\$my*. *To* is the object of the subordinate clause and is generated to the right of the verb. Consider:

- (6) *My zrobili\$my to* 'We did this'
 [IP ...[V" [N" My] [V' [V zrobili\$my][N" to]]]]

The diagram represents a simplified structure of the clause in 6.⁵ In Polish, objects are free to move within a clause. This does not apply only to complements, but also other elements, hence Polish is a very free word order language.⁶ Going back to our example in (5b), we might argue that the preterite endings either cannot attach to moved material, or that cliticization precedes object movement. It turns out, however, that both hypotheses are wrong. Consider the following example:

- (7) (a) *On wiedział, że my jemu+\$my kupili książkę*
 he know+past that we him+CL buy+past+gender book
 'he knew that we bought him a book'
 ?(b) *On wiedział, że my jemu książkę+\$my kupili*
 *(c) *On wiedział, że my kupili jemu+\$my książkę*

As we can see from (7a), the clitic can attach to the indirect object. It seems also possible for it to attach to the direct object, although it sounds rather odd (7b). (7c) shows that the clitic cannot attach to the same object before it has moved in front of the verb. From this we can assume, that for the preterite endings (this holds for *by* as well) any material which is linearly placed after the verb cannot serve as a host for the clitics.

On the whole, both the selectionally independent *by* and the preterite clitics are free to attach to anything before the verb (with minor exceptions). However, this restriction applies only to the linear configuration after movement has taken place. As we have seen, objects originating within the VP and placed to the right of the head verb can serve as hosts after they have been fronted.

This requires that Polish lexical rules, such as Lower and Raising (see section below) be allowed to operate in the postlexical component provided they are restricted in their operation to the domain of phonological words.

4. The Phonology of Polish Clitics

The most relevant observation about clitics, which also concerns affixes, is that they tend to form one phonological unit with an independent form. Such a unit is usually called *mot* or a phonological word.⁷

⁵ I omit the AGR nodes, as well as the Tense and Negative Phrases and their specifiers. For clarity's sake also, I did not indicate the Det phrases.

⁶ See Willim 1989.

⁷ I assume the definition of a phonological word, and a prosodic hierarchy as adopted by Nespor and Vogel 1986.

A typical characteristic of clitics, differentiating them from separate words, is their accentual dependency (Zwicky (1985:287)). In Polish the attachment of preterite endings causes stress shift for the majority of native speakers.⁸ Consider the following data concerning *\$my*:

- (8) (a) *On załatw+i+l Janka* 'he got away with John'
 (b) *Oni załatw+i+l+i Janka* 'they got away with John'
 (c) *My załatw+i+l+i+\$my Janka* 'we got away with John'

In (8a-b) the stress on the verb is on the penultimate syllable, in (8c) the stress may be shifted one syllable to the right. This indicates two things: the preterite endings do not carry stress of their own and, after their attachment, re syllabification must be allowed to apply.

Let us consider the Polish 1st. person pl. preterite clitic [*\$mĩ*]. When it is attached to the verb we obtain *poszli\$my* /pošlišmĩ/ 'we went'. There is a tendency, especially among younger speakers, to place the stress on the second syllable. However, many speakers place stress on the first syllable, as in the verb *poszli* /pošli/ 'they went'. In the latter case, there is no stress distinction between these two forms. This implies that, although there is a tendency to shift the stress after the clitic has been attached, it is not a regular phenomenon.⁹

The rule which imposes stress must take place in the word domain, because in phrases like: *poszli daleko* 'they went far' the stress is on the first syllable of the verb, and cannot be changed even by innovative speakers. This is because *poszli daleko* is not a phonological word, but two independent mots. Consider¹⁰:

- (9) [mot [s poš][s li]] [mot [s dal][s le] [s ko]]

In the case of *poszliście* /pošlišće/ 'you went' ([šće] being the 2nd. person plural preterite ending), or in *poszli\$my* /pošlišmĩ/ 'we went' ([šmĩ] being 2nd. person plural preterite ending), the construction forms a single phonological word.

- (10) (a) [mot [s poš][s li] [s \$mĩ]] (b) [mot [s poš][s li] [s šće]]

In both cases (10a, b) the stress, for many speakers, may be shifted to the second syllable.

Let us presently consider Polish Raising,¹¹ a rule which turns /o/ into /u/, for instance in: *Bog+a* /boga/ 'God' (gen.sg) vs. *Bóg* /buk/ 'God' (nom. sg.). The context for raising is a non-syllabic voiced consonant following a vowel. Consider the verb *mogłem* [[[mog]w]e]m] /mogwem/ 'I could', and compare it with *mógł* [[mog]w] /mugw/ 'he could'. If the floating 1st. person sg. preterite clitic *-m* is attached to the verb, then Raising does not occur. The same situation arises with the 2nd. person sg., 1st. person pl., 2nd. person pl. preterite clitics: [*\$*], [*\$mĩ*], [*\$ście*]. Consider the following examples:

- (11) *mógł* [[mog]w] /mugw/ 'he could' vs. *mogłeś* [[[mog]w]eš]] /mogweš/ 'you could', *możliście* [[[mog]l]i]šće] /możliście/ 'you could (pl)', *mogłem* [[[mog]w]em] /mogwem/ 'I could.'

From (11) we can see that the preterite clitics create a context for the application of lexical phonological rule - they block Raising just like affixes, and hence they behave phonologically no differently than affixes (see Booij & Rubach (1987), Szczegielniak (1991), and Zwicky (1977) and (1985)).

Rubach (1984:184ff.), proposes that certain e's alternate with Ø and/or with i/ɨ (the latter spelled y) and some do not. Polish e's alternating with zero and/or i/ɨ derive from underlying high lax vowels /i̯ ɨ̯/ one of which is palatalising and the other not. These

⁸Some older speakers do not shift stress after a preterite ending is added.

⁹Booij & Rubach (1987) claim that stress shift is only obligatory when the 1sg. and 2sg. preterite clitics are attached.

¹⁰At this stage of the derivation morphological boundaries are erased. However, for the sake of clarity, I will keep specifying them whenever necessary, either by using square brackets or the plus sign.

¹¹I suggest a statement of the rule along the lines proposed by Booij & Rubach (1987:10):

$$O \rightarrow U / __ C_0 \left[\begin{array}{c} -syl \\ +voic \end{array} \right] \quad]]$$

RAISING

The double square brackets indicate the end of a constituent (a phonological or morphological word).

vowels, called *yers*, either lower to /e/ when followed by a yer in the next syllable, or delete context-free.

The process responsible for the alternations is Lower: is a lexical rule applying cyclically. It interacts with Yer Deletion which also applies in the lexicon, but it is not cyclic.¹² The underlying representation of *mógł* [[mug]w] //mug+ɫ+ɨ/ 'he could', has to be revived, following Rubach (1984), I assume that the final yer in the Underlying representation of *mógł* is a gender morpheme (see the derivation of *szedł* below). Hence, Raising must apply after Yer Deletion otherwise the final yer will block the application of the rule. Booij & Rubach (1987:11) argue that both Yer Deletion and RaisingLower are Postcyclic rules, that is rules applying lexically but not cyclically. This means that the context for Raising is the end of a word and not some other constituent.

I have already shown that preterite clitics can block the application of Raising, let us now have a look at how they interact with Lower. Booij & Rubach (1987:37) explain the alternation between *szedłem* 'I went' and *szedł* 'he went' by proposing the derivation below:

| | | | |
|-------------|----------------|--------------|--------------------|
| (12) | <i>szedłem</i> | <i>szedł</i> | |
| Cycle 1 | šĩ d | šĩ d | |
| Cycle 2 | šĩ d+ɫ | šĩ d+ɫ | WFR: Preterite |
| Cycle 3 | šĩ d+ɫ+ɨ | šĩ d+ɫ+ɨ | WFR: Gender |
| | šed+ɫ+ɨ | šed+ɫ+ɨ | Lower |
| Cycle 4 | šed+ɫ+ɨ+mĩ | ----- | WFR: Cliticization |
| | šed+ɫ+e+mĩ | ----- | Lower |
| Postcyclic: | šedwem | šedw | |

As we can see in (12), the presence of the preterite clitics triggers of the application of Lower (the above derivation applies to all the other preterite endings). Note that I assume that the Underlying Representation of *ɫ* and of *m* is respectively //šĩ //, //mĩ // (for a detailed argumentation for adopting such a representation see Booij & Rubach (1987)).

Polish *by* 'would' does not exhibit the phonological and morphological properties of the preterite endings. It does not induce stress shift, nor does it trigger the application of phonological lexical rules, like Raising, or Lower. Consider:

- (13) (a) *On mógłby* /mugwbɨ/ *pobiec* 'he could have run, if...(3rd. conditional)'.
 (b) *My mogliśmy* /moglišmɨ/ *pobiec* 'we could have ran, but we did not'
 (c) *On szedłby* 'he would have went'

Szczegielniak (1991:30) points out that Polish *by* is somewhat different from the preterite endings. It exhibits clitic properties in the syntax, but phonologically it behaves like a separate word, and in that sense it is different. Consequently, we must conclude that *by* and the preterite clitics constitute two separate sets of clitics.

The preterite endings are special clitics, in the sense that they do not have corresponding full forms; Zwicky (1977:8) distinguishes special clitics from simple ones. An example of a simple clitic is the reduced form of the future tense marker 'will'. Simple clitics, on the other hand, usually occupy the same positions as their full counterparts.¹³ The problem is that the distinction between simple and special clitics adopted by Zwicky (1977:8) is not sufficient for Polish. Both the preterite clitics and *by* are special clitics. In

¹² I will adopt the following statement of Lower and Yer Deletion (Booij & Rubach 1987:9):

Lower (I use the symbol /ɫ/ to denote /ɨ/):

Yer Deletion: $\left\{ \begin{array}{c} \text{ĩ} \\ \text{ĩ} \end{array} \right\} \rightarrow e / \text{---} C_0 \quad \left\{ \begin{array}{c} \text{ĩ} \\ \text{ĩ} \end{array} \right\} \rightarrow \emptyset$

¹³ P.H. Matthews (personal communication) has rightly pointed out to me that the criterion of position does not mean that a clitic has a full corresponding form. For example, full forms of pronouns do not share the same distribution with their clitic counterparts. This is true of Polish (see Aguado & Dogil (1989)). Conversely, Latin 'and/or' do not share the same positions as other co-ordinators, and have no full forms.

order to distinguish the two types of clitics, I will postulate that *by* is a word-like clitic, since it differs from separate words only in its syntactic behaviour. The preterite endings, on the other hand, are affix-like clitics, since they differ from independent words both syntactically and phonologically.

Another observation concerning the syntax of clitics, is the fact that they can serve as hosts to other clitics, just like affixes (see Zwicky and Pullum (1983:504)), see examples (5f,k-m). In fact, the preterite endings must attach to the *by* clitic, forming the following construction:

- (14) [CLITIC1 [CLITIC1 *by*] [CLITIC2 Preterite ending]].

Following Szczegielniak (1991), I will claim that Polish *by* and Polish preterite endings are a form of phrasal affixation, and hence, should trigger the application of Lexical rules just as affixation taking place in the Lexicon does. The question to ask is what is the status of the relevant processes. If we assume that Polish preterite endings are generated in the syntax then Lower, Raising, Yer Deletion and Stress Assignment cannot be restricted in their application to the Lexicon. Let us assume that the above mentioned processes are restricted in their application in a different manner than that proposed by Booij & Rubach 1987.

I will adopt a model proposed by K. Rice (1990) where it is proposed that the domains of rule application in the lexical phonology are also relevant at phrase level phonology. Phonological rules are available from the beginning of phonology, however, their application is governed by three interacting principles:

- (15) (a) **STRONG DOMAIN HYPOTHESIS (SDH)**: The grammar may stipulate merely where a rule ceases to apply. All rules are potentially applicable at the first level of the lexicon, and apply provided only that the principles of the grammar permit it; at lower levels of the lexicon, and in the postlexical phonology, rules may be 'turned off' but no new ones may be added (Kiparsky 1984).
 (b) **STRICT CYCLE CONDITION (SCC)**: If *W* is derived from a lexical entry *W'*, where *W'* is non distinct from *XPAQY* and distinct from *XPBQY*, then rule *A->B/ XP_QY* cannot apply to *W* until the word level (Kiparsky (1985:89)).
 (b) **STRUCTURE PRESERVATION (SP)**: Non distinctive features and structures cannot be introduced in the lexical phonology (following Kiparsky 1985).

Following SDH all rules are available from the beginning of the derivation until they switch off. In reality not all rules apply at the first available domain. This, argues Kiparsky, is caused by the interaction of SCC and SP.

Selkirk (1980) and Vogel (1984) propose the existence of a set of rules applying only in given domains: **domain span** rules. The set includes sandhi rules that apply across words, phonological phrases, intonational phrases and utterances. Vogel (1984) distinguishes domain span rules which are allophonic in nature, for example flapping in English. Domain span rules applying to smaller domains, on the other hand, are not allophonic. Vogel (1984) gives examples of Italian S-Voicing and the American English Rhythm Rule. Hence, domain span rules applying at word and phonological phrase level have either neutralising or structure building properties of lexical rules.

Apart from domain span rules, Vogel (1984) proposes that there is a set of **domain limit** rules that apply at word edges and at pauses. Vogel identifies several word edge rules: vowel tensing in American English and depalatalization in Sanskrit. These rules are neutralising. For example, American English vowel tensing tenses word-final vowels at the edge of words, taking effect when a word is the first member of a compound and when it stands on its own.

Let us assume that Lower is a domain span rule. It applies in the domain of phonological word, and is cyclic (constrained by the SCC). There is no contradiction in a rule being cyclic and domain sensitive. As Goldsmith (1990:25) puts it, "the notion of cyclic derivations derives from the idea that certain morphological processes may themselves take as their input certain objects that have already been turned into well formed

words. Certain generalisations that involve word level units may then hold both of the larger unit and the smaller unit, but ultimately this should be no more surprising or controversial than the proposition that main clauses may contain subordinate clauses.

As I have shown on the example of Polish stress, affixation in the lexicon means that the affix is incorporated into the prosodic structure of the phonological word. Therefore there is overlap between the statement that Lower applies in every cycle and that it applies after the mot structure of a given element has altered. It seems that Booij & Rubach (1987) assume that only processes in the lexicon can cause a change in the prosodic structure of a word. However, if we assume that affix-like cliticization can also change the structure of mot, then Lower should also be allowed to apply. This means that Lower is 'turned off' after the last rule influencing the structure of mot has applied, this does not necessarily mean it has to take place in the lexicon. Let's assume, YER deletion also has the status of a domain sensitive rule. It differs from Lower in that it is not bound by the SSC, since it is context-free.

Finally let us consider the status of Raising. As Booij & Rubach (1987) point out, one of the contexts for Raising is that it must apply at the end of a word. I propose that Raising is a domain limit rule and that its application is restricted to the edge of mot. This is a stronger prediction than that of Booij & Rubach's, since they do not specify what exactly they understand to be a word. In the case of Polish cliticization and its interaction with phonological rules, the relevant domain is a phonological word. Notice that *by* can be incorporated morphologically to the verb, however, it does not trigger any phonological processes.

At first glance we seem to lose the generalisation concerning the ordering of the rules above. The fact that Lower precedes YER deletion and Raising is predictable from the nature of the rules. Lower, according to Booij and Rubach, is a cyclic lexical rule, YER deletion is Postcyclic and lexical, and so is Raising. However, if we assume that the application of rules is governed by the three principles in (15), we can impose certain ordering restrictions. Lower is constrained by the SCC and is structure preserving. YER deletion is in fact a reflex of a more general principle *Stray Erasure* which deletes at the end of a derivation material which is not prosodically adjoined to mot (see Steriade (1982), Booij & Rubach (1990)). Yers which are not vocalised cannot be assigned a syllable structure, and hence cannot be incorporated to mot.¹⁴ Stray erasure applies at the end of a given derivation, for example after the application of domain sensitive rules. Raising, being a domain limit rule, will apply after all the domain sensitive rules have applied. I am not postulating that domain sensitive rules have to precede domain limit ones, however, a domain sensitive rule constrained by the SCC and structure preserving, has to precede a neutralising domain limit rule not constrained by the SCC.

5. Cliticization and Wh-Movement in Polish

The distribution of *by* and the preterite endings is more restricted than that of independent words, but whether clitics undergo movement is another matter. In one of the sections above I showed that cliticization applies after the object has moved in front of the verb. Let us now consider the interaction of Wh movement with cliticization in Polish. Polish, like English, has obligatory Wh-movement. Consider the following examples:

- (16) (a) *Co on wiedział* 'what did he know'
 What he know+past
 (b) *On wiedział có*
 (c) *On có wiedział*

¹⁴Following Kenstowicz & Rubach (1987), I assume that yers are floating matrices, and cannot play any role in syllable structure

Cases (16a-c) indicate that Wh phrases in Polish do have to move like in English. (16b) is identical to the English *He knew what* with emphasis on the last word. Case (16c) is a result of the fact that in Polish verb complements can be moved in front of the verb. However, the Wh word still has to have emphatic stress on it, and hence it is no different in status from case (16b). Following Chomsky (1991) and Lasnik & Saito (1992), I assume that Wh-movement also takes place at LF. I will assume, following Lasnik & Saito (1992:1-17), that Wh movement is constrained the syntax by Subjacency, and that the ECP constrains Wh movement in the syntax and LF¹⁵. Following Lasnik & Saito (1992:56), I will assume the existence of filters governing Wh-movement in syntax and LF.

- (17) (a) A COMP is [+WH] iff a [+WH] phrase is in its SPEC position.
(b) A head and its SPEC cannot be contraindexed.

This implies that there is no level of grammatical representation where Wh movement is not constrained, although it has more freedom in LF.¹⁶

Let us look at some examples of Wh movement interactions with cliticization in Polish.

- (18) (a) *Co on wiedział, że my zrobiliśmy* 'what did he know we did'
[_S [_{CP} Co [_S on wiedział [_{CP} t₁ że [_S my zrobiliśmy t₁]]]]]

The trace t₁ is necessary, since otherwise Subjacency would be violated (two barriers are crossed the NP headed by *my* and CP containing *że*). Notice that t₁ is moved to the SPEC-COMP position of the embedded clause, so is the Wh phrase at the beginning of the clause.

It seems that in Polish Wh clauses cannot be Wh moved if they are a clitic host.

Consider:

- (19) (a) *On wiedział, że cośmy' my zrobili* 'he knew that we did what'
[_S [_{CP} [_S on wiedział [_{CP} że [_S co+sm' my my zrobili t₁]]]]]
*(b) *Cośmy on wiedział, że zrobili*
[_S [_{CP} Cośmy [_S on wiedział [_{CP} t₁ że [_S my zrobili t₁]]]]]

One might argue that the trace t₁' in (19b) cannot antecedent govern t₁ and thus the ungrammaticality of (19b) is a result of violating the ECP and has nothing to do with cliticization. Notice that in (18a), the embedded trace of the Wh clause is also not antecedent governed by the trace in COMP, and yet the sentence is perfectly grammatical.

In order to resolve what is directly responsible for the ungrammaticality of (19b), consider the following sentences:

- (20) (a) *Kto on podejrzewał zabił Janka* 'who did he suspect killed John'
who he suspect+past kill+past john+dat
[_S [_{CP} Kto [_S on podejrzewał [_{CP} t₁ [_S t₁ zabił Janka]]]]]
(b) *On podejrzewał, że kto zabił Janka*
he suspect+past that who kill+past john(dat)
[_S [_{CP} [_S on podejrzewał [_{CP} że [_S kto zabił Janka]]]]]
*(c) *Kto on podejrzewał, że zabił Janka*
who he suspect+past that kill+past john+dat
[_S [_{CP} Kto [_S on podejrzewał [_{CP} t₁ że [_S t₁ zabił Janka]]]]]

The above examples indicate that in the case of Wh subject traces, the ECP is violated and t₁ is left ungoverned. Lasnik & Saito (1992) point out, that, in cases like (20c), the fact that

¹⁵ECP (Empty Category Principle):

1. A nonpronominal category must be properly governed (Chomsky 1981)
2. a properly governs b iff a governs b and
 - a. a is a lexical category X0 (lexical government), or
 - b. a is coindexed with b (antecedent government)
3. a governs b iff every maximal projection dominating a also dominates b and conversely (Aoun & Sportiche 1982/83). Lasnik & Saito (1992) propose a modified account of ECP, however, for present purposes the above definition will suffice.

As far as Subjacency is concerned, I will adopt a definition proposed by Chomsky (1986).

¹⁶ For an account of Wh-movement in Polish within the GB framework see Willim 1989

the head of COMP is *ze* carrying the feature [-wh] (since it is not a Wh word), t_1' (occupying the SPEC of COMP) cannot govern the subject trace. This is because the index of the SPEC is copied onto the head only if the SPEC and the head agree with respect to the feature [+WH]. In the case of (20a), the trace t_1' situated in the SPEC of COMP position can index the head, since there is no [-WH] element occupying that position. Once the head of COMP is coindexed with t_1' it can antecedent govern t_1 . In the case of object traces the situation is different: the Wh trace is not antecedent governed but is governed by the verb. Hence, the ECP is not violated and (18a) is grammatical. Returning to our example in (19b), I will assume that cliticization is a form of head-head movement and the preterite endings are moved from AGRS and cliticized to a given host, as in (19a). In general, Wh movement cannot interact with cliticization in Polish. The above analysis predicts that in Polish Wh-moved phrases cannot serve as clitic hosts.

Let us, once more, have a look at example (19a), where the clitic is attached to a Wh word and the sentence is perfectly grammatical. The fact that *co* did not undergo Wh-movement is not enough to explain why it can serve as a clitic host, since it does occupy the SPEC-AGRO position. One possible answer is that Wh-movement precedes head-head movement, and hence *co* in (18a) moves to SPEC-AGRO in order to obtain case, but only after it finishes moving to the SPEC of the matrix COMP can cliticization take place. Notice that in cases like (19a) *co* has only moved¹⁷ up to the SPEC-AGRO of the embedded clause, and is a valid host for the clitic. The general conclusion is that, for Polish, cliticization of the preterite endings via head-head movement cannot cross to the left of COMP and to the right of the head of V". This is a reasonable prediction, since the preterite clitics are agreement markers within the clause, and hence should not be able to undergo clitic climbing as for instance cliticized pronouns in French do (see Roberts 1993b).

It is essential to note, however, that I am not postulating that movement to SPEC positions generally precedes head-head movement (regardless whether it is optional or obligatory). I mainly assume that T obligatorily raises to AGRS in overt syntax, because in Polish tense and person/number agreement are visible in the same paradigm. Hence, the verb must be able to check those features 'in one go'.

6. Cliticization and Infinitival Clauses in Polish

Let us consider the case when TP is [-finite]:

- | | | | |
|------|-----|------------------------------|---------------------------------|
| (21) | (a) | <i>On chciałby go zabić</i> | 'he would like to kill him' |
| | | he1 like+past+CL him2 kill | |
| | (b) | <i>On chciał go by zabić</i> | 'he wanted someone to kill him' |
| | | he1 want him2 CL kill | |
| | (c) | <i>On chciał by go zabić</i> | 'he wanted someone to kill him' |
| | | he1 want him2 CL kill | |
| | | he1 want him1 CL kill | |

In examples like (21b-c), *by* cannot be in the COMP position since then it would govern PRO in the SPEC-AGRS position of the subordinate clause:

- (22) [_C [_C [_C *by*]_{AGRS} PRO [_{AGRS} [_{AGRS} T{-finite}]]]]

A plausible solution is that verbs like *chcieć*, taking infinitival clauses as complements, do not select *by* in COMP, and the clitic is generated in the T [-finite] of the subordinate clause. This is supported by the semantic differences of the sentences in 21. The meaning of the clause when *by* is attached to the main verb is different from the interpretation assigned to a construction when *by* is generated in the subordinate clause.

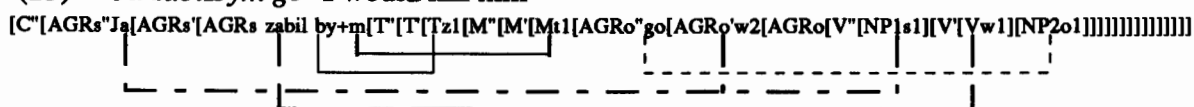
The above data indicates that *by* is independent of the V features in T. We can also assume that since, the preterite clitics cannot appear when T is [-finite], there must be some

¹⁷ Note that I do not specify whether this is obligatory or optional movement.

restriction on their generation in AGRS. One plausible explanation is that, when the verb is [-finite] the subject of the clause is PRO with which there is no subject-verb agreement.

The independence of *by* from the features in T indicates that *by* cannot be the head of TP. G. Horrocks (personal communication) points out that *by* is a mood and not tense marker, and hence should rather be a head of a mood phrase. It still would have scope over the whole phrase, since it would c-command it. Notice also that the head of MP has to raise to AGRs before the head of TP is raised to AGRs, otherwise we would not be able to generate the order tense+*by* + preterite clitic. If we assume that T raises in overt syntax, then we must also postulate the overt raising of M. Consider the following example:

(23) *Ja zabiłbym go* 'I would kill him'



It is not important for the analysis here whether the relative movements are obligatory or optional. I only assume that T and M raise obligatorily in overt syntax. Notice that the feature matrix in AGRS reflects the morphological distribution of the listed elements: *Ja zabił*+t (tense)+ *by* (Clitic1) + *m* (Clitic2) *go*. The only necessary condition is that the clitic has a valid host to which it can cliticize.

Summing it up, we can postulate that the preterite endings are generated in AGRS and later cliticize obligatorily to *by* (if it is present) which is moved up from the head position of the mood phrase MP. If *by* is not present, the preterite clitics cliticize to any element between the head of COMP and the head of VP.

The distribution of *by* depends on the selectional properties of the main clause verb. Verbs like *chcieć* 'to want' select *by* in COMP, if the complement clause is [+finite]. Verbs like *wiedzieć* 'to know' do not select *by* in COMP, and it is generated in the head of MP of the subordinate phrase. The situation is identical in the case of [-finite] clause complements of the verbs like *chcieć*.

7. The Structure of Cliticization in Polish

Let us consider what are the exact processes concerned with cliticization in Polish. The clitics themselves are listed in the lexicon and have a [+cl] feature assigned to them. However, neither affix nor word-like clitics are visible to lexical processes; they cannot be introduced into the derivation by Word Formation Rules applying in the lexicon. The preterite clitics are generated in the syntax under AGRS0, since they are in fact, a phonetic manifestation of the verb-subject agreement. *By* is generated under M, or in COMP depending on the main clause verb. The nodes under which the clitics are generated are, in fact projections of their syntactic host (the verb). *By* obligatorily moves up and incorporates to AGRS0, hence, if both clitics are generated, the preterite clitics must be adjoined to *by*.¹⁸ Both the *by*+preterite clitic construction and the clitics themselves can attach to lexical elements between the head of COMP" and the head of V". This is done through movement, but movement to a position occupied by a lexical head. Both word-like and affix-like clitics cannot be left in a clause initial position. This is a reflex of the Tobler-Mussafia Law, which can be interpreted as a requirement that something precedes AGR10 (here AGRS), which a manifestation of the ECP.¹⁹ In Polish, Cliticization is preceded by Wh-movement and does not interact with it. Cliticization is not only a syntactic process involving movement. Affix-like clitics are further incorporated into the prosodic structure of their phonological host. This triggers off the application of domain span and domain

¹⁸ This has to be feature driven movement

¹⁹ From Roberts (1993a).

limit rules. As we can see, Word-like cliticization is a purely syntactic process, whereas affix-like cliticization is a syntactic and later phonological process.

Let us consider return to Polish stress assignment. Rubach & Booij claim that the clitics *ście* and *my*, for some speakers can still not trigger stress shift. This can be explained by the model adopted above, since 1 sg. preterite clitic *m* and 2sg. preterite clitic *s* do not have a prosodic structure of their own, since they do not have a syllable nucleus.²⁰ This means that *m* and *s* do not have to undergo mot deletion, because there is nothing to delete, and they have to be incorporated into the prosodic structure of their host independently of cliticization in order not to be deleted by Stray Erasure. Hence, we can postulate that stress assignment, for some speakers, takes place before cliticization but after independently motivated prosodic incorporation. This would explain why the plural preterite clitics do not cause a stress shift for some speakers. The other possibility is that stress assignment follows cliticization and, as a result, all the preterite clitics cause stress shift. To sum it up, both word-like and affix-like clitics are adjoined in the syntax via movement. Affix-like clitics are then incorporated prosodically into their host, either because they have no syllable structure of their own, as in the case of the singular preterite clitics, or through mot deletion and adjunction as in the case of the plural preterite clitics. The above analysis has certain advantages compared to the assumption that cliticization takes place in the lexicon, since Booij & Rubach (1987) would have to stipulate that the plural preterite clitics are some sort of an exception to stress assignment.

8. Interpretation of Cliticization at LF

Let us consider a situation when there is no obligatory movement at all²¹, only T raises to AGRS; *by* and the preterite clitic are situated in AGRS0. In LF where the effects of optional movement are not visible, the verb is the AGRS0 position and both NP's are located in SPEC of AGRS, and in SPEC of AGRO respectively.²² The verb is raised to the clitic complex situated in AGRS. This implies that a situation when both clitics are attached to the verb is the unmarked case in Polish. This is in line with the data concerning word-like and affix-like clitics. Cases when either of the clitics is attached to an element different from the verb are considered to be marked.

Chomsky (1993) assumes that overt syntactic movement is 'more expensive' than movement in LF. In Polish, the optional overt movement of clitics creates marked constructions. The meaning of the clause is identical to the meaning of a clause in which all the movements are carried out at LF. However, there is a difference in emphasis. The element to which the clitic attaches (apart from the verb) receives additional emphasis. I assume that the motivation for the clitics to move overtly is pragmatic - to emphasise certain elements in the clause.

9. Conclusions

I have tried in this paper to give a concise account of the cliticization phenomena in Polish. I claim that cliticization is a form of phrasal affixation.²³ The discussed clitics are generated in syntax and their marked distribution (other than adjoined to the verb) is a result of the

²⁰ Polish does not have syllabic consonants.

²¹ Of course, something has to move optionally otherwise the clitic would not have a valid host. It is essential to note that optional movement does not form a chain. See Poole (to appear).

²² J. Bobaljik & D. Jonas (1993) claim that in languages with tense and agreement morphology in the same inflectional paradigm license two SPEC positions: SPEC AGRS and SPEC T ". However, for the purpose of the analysis herein I think that this does not have any significance.

²³ This is in line with the claims made in Anderson (1992)

application of overt syntactic head-head movement. In order to describe the facts, I assume that Polish free word order is a result of the optionality of overt head movement. In order to account for the fact that Polish lexical rules interact with cliticization, which is a syntactic process, I put forward a model in which phonological rules in Polish apply in certain domain defined environments.

The model presented in this paper does not give a full account of cliticization in Polish. For one, it only discusses certain clitics. A strong drawback of this analysis is the fact Polish syntax still has not been described adequately. Therefore, I was forced to assume certain analysis of Polish syntax which upon further research might prove inadequate. I have also not discussed in detail the syntactic processes connected with cliticization (how the clitics are adjoined, etc.).

Submitted: April 1995

References:

- Aguado, M., and G. Dogil. 1989. Clitics in lexical phonology: alleged counter evidence? *Linguistische Berichte* 120.
- Anderson, S.R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Aoun, J., and D. Sportiche. 1982/83. On the formal theory of government. *The Linguistic Review* 2, 211-236.
- Bobaljik, J. D., and D. E. Jonas. 1993. Subject positions and the role of TP. Paper presented at the 1993 GLOW meeting, Lund.
- Booij, G., and J. Rubach. 1987. Postcyclic versus postlexical rules in lexical phonology. *Linguistic Inquiry* 18 :11-44.
- 1990. Edge of constituent effects in Polish. *Natural Language and Linguistic Theory* 8, 427-463.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Foris, Dordrecht.
- 1986. *Barriers*. MIT Press, Cambridge, Mass.
- 1991. Some notes on the economy of derivation and representation. In *Principles and Parameters in Comparative Grammar*. MIT Press, Cambridge, Mass.
- 1993. A minimalist programme for linguistic theory. In *The view from Building 20: Essays in linguistics in honor of Sylvian Bromberger*, ed. Kenneth Hale & Samuel Jay Keyser, Cambridge Mass.: MIT Press, 1-52.
- Dogil, G. 1984. Lexical phonology and floating inflection in Polish. In W.U. Dressler (ed.), *Phonologica* 1987, Cambridge: Cambridge University Press.
- Goldsmith, J.A. 1990 *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.
- Hale, K. 1973. Person marking in Walbiri. In S.R. Anderson and P. Kiparsky (eds.), *Festschrift for Morris Halle*, Holt Reinhart and Winston, New York, pp. 308-344.
- 1985. On nonconfigurational structures. Manuscript, MIT.
- Halle, M., and K.P. Mohanan. 1985. Segmental phonology of Modern English. *Linguistic Inquiry* 4,3-16.
- Hayes, B. 1990. Precompiled phrasal phonology. In S. Inkelas and D. Zec (eds.), *The Phonology Syntax Connection*. Chicago: The University of Chicago Press.
- Hoeksema, J., and R. D. Janda 1988. Implications of process morphology for categorial grammar. In R. T. Oehrle, E. Bach, and D.Wheeler (eds.), *Categorial Grammars and Natural Language Structures*, 199-247. Dordrecht Reidel.
- Kenstowicz, M.& J. Rubach. 1987. The phonology of syllable nuclei in Slovak. *Language* 63: 463-97
- Kiparsky, P. 1982a. Lexical morphology and phonology. In I.-S. Yang (ed.), *Linguistics in the Morning Calm*. Seoul: Hanshin.
- 1982b. From cyclic phonology to lexical phonology. In van der Hulst and Smith 1982, Part 1.

- 1984. On the lexical phonology of Icelandic. In C.-C. Elert et al. (eds.), *Nordic Prosody III: Papers from a Symposium*, 135-162. University of Umea.
- 1985. Some consequences of lexical phonology. *Phonology Yearbook* 2, 85-138.
- Klavans, J. L. 1980. *Some Problems in a Theory of Clitics*. Bloomington: Indiana University Linguistics Club.
- 1985. The independence of syntax and phonology in cliticization. *Language* 61, 95-120.
- Klemensiewicz, Z., T. Lehr-Splawiński, and S. Urbanić. 1955. *Gramatyka historyczna języka polskiego*, Warsaw: Państwowe Wydawnictwo Naukowe.
- Lasnik, H., and M. Saito. 1992. *Move α : Conditions on Its Application and Output*. MIT Press, Cambridge, Mass.
- Matthews, P. H. 1972. *Inflectional Morphology: a Theoretical Study Based on Aspects of Latin Verb Conjugation*. Cambridge: Cambridge University Press.
- McCarthy, J. J., and A. S. Prince. 1990. Foot and word in prosodic morphology; the Arabic broken plural. *Natural Language and Linguistic Theory* 8, 209-263.
- McHugh, B. D. 1990. The phrasal cycle in Kivunjo Chaga tonology. In S. Inkelas and D. Zec (eds.), *The Phonology-Syntax Connection*. Chicago: The University of Chicago Press.
- Nespor, M., and I. Vogel. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- Nevis, J. A. 1985. Finnish particle clitics and general clitic theory. Ph.D. thesis, Ohio State University.
- Poole, G. Optional Movement in the Minimalist Programme. To appear in *Minimal Ideas*, W. Abraham, S.D. Epstein, H. Thrainsson, C.J. Zwart (eds.) Benjamins: Philadelphia.
- Pollock, J-Y. 1989. Verb movement, Universal Grammar and the structure of IP. *Linguistic Inquiry* 20, 365-424.
- Rice, K. D. 1990. Predicting rule domains in the phrasal phonology. In S. Inkelas and D. Zec (eds.), *The Phonology-Syntax Connection*. Chicago: The University of Chicago Press.
- Roberts, I. 1993a. Aspects of clitic-placement in Old Germanic and Old Romance. Paper presented at the 1993 GLOW Workshop, Lund.
- 1993b. Clitic dependencies and A-dependencies. Paper presented at the 1993 GLOW meeting, Lund.
- Rubach, J. 1984. *Cyclic and Lexical Phonology: The Structure of Polish*. Dordrecht: Foris.
- Selkirk, E. O. 1980. Prosodic domains in phonology: Sanskrit revisited. In M. Aronoff and M.-L. Kean (eds.), *Juncture (Studia linguistica et philologica 7)*, 107-129. Saratoga, Calif: Anna Libri.
- Spencer, A. 1991. *Morphological Theory*. Cambridge University Press.
- Steriade, D. 1982. Greek prosodies and the nature of syllabification. Ph.D. dissertation, MIT.
- Szczegielniak, A. 1991. The problems caused by Polish clitics in the model of lexical phonology. *Bulletin de La Societe Polonaise de Linguistique*, fasc. XLVI, 1991, 30-39.
- 1993. Cliticization in Polish. Unpublished M.Phil dissertation. Cambridge University.
- 1995. Economy of Movement, the Nature of Features, Move α , and Form Chain--the Case of Polish. Poster presented at 16 GLOW Colloquium, Tromsø.
- Vogel, I. 1984. On constraining prosodic rules. In H. van der Hulst and N. Smith (eds.), *Advances in Nonlinear Phonology*, 27-233. Dordrecht: Foris.
- William, E. 1989. *On Word Order: a Government-Binding Study of English and Polish*. Zeszyty Naukowe Uniwersytetu Jagiellońskiego, zeszyt 100.
- Zwicky, A. M. 1977. *On Clitics*. Bloomington: Indiana University Linguistics Club.
- 1985. Clitics and Particles. *Language* 61, 283-306.

Zwicky, A. M., and G.K. Pullum. 1983. Cliticization vs. Inflection: English n't.
Language 59, 502 -513.

Symbols:

| | |
|---|--|
| /// // [] CL S š, ž č ś, ź î | Underlying representation Phonological representation Morphological representation Clitic Syllable Postalveolar fricatives Prepalatal affricates Prepalatal fricatives Non-palatalising high front vowel, IPA symbol /i/ |
|---|--|